

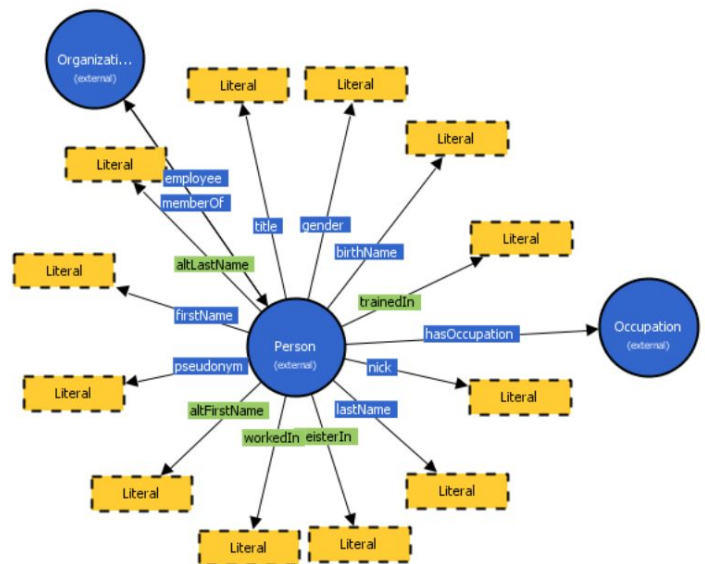
## *Cultural Heritage Data Extraction and Knowledge Graph Construction*

### Objective of this work:

Cultural heritage institutions store and digitize large amounts of data inside archives to make archival records findable by archivists, scientists, and general public. The important heritage value for both socio-historical work as well as family and genealogical research present person registers, which contain personal information, e.g. family relations, names, occupation, place of residence, birth or death date, etc. Knowledge Graphs can be used to bring these historical data together and make them findable and usable for users.

In this work, the challenges of structuring historical data and KG construction will be investigated. The use case of the work are 20 address books with name, address, profession and other records of persons, who resided in Nürnberg from 1890 till 1910 (the data is in German). The resources are already available in a textual digital form (OCR).

As a first step of the work, the data obtained from the address books will be transformed to semi-structured data using existing tools for text analysis, e.g. regular expressions. The identified entities as well as their relations and classes will be integrated within the Knowledge Graph to be explored. Moreover, to enrich the knowledge base, the entities will be connected to external sources, e.g. Wikidata [1] and GND [2].



[1] <https://www.wikidata.org/>

[2] [https://www.dnb.de/EN/Professionell/Standardisierung/GND/gnd\\_node.html](https://www.dnb.de/EN/Professionell/Standardisierung/GND/gnd_node.html)

The project work will be supervised by **Prof. Dr. Harald Sack, Tabea Tietz and Oleksandra Vsesviatska**, Information Service Engineering at Institute AIFB, KIT, in collaboration with FIZ Karlsruhe.

### Keywords:

Knowledge Graphs, Cultural Heritage

### Pre-requisites:

Knowledge of Programming with Python.

Contact persons:

**Tabea Tietz**

tabea.tietz@fiz-karlsruhe.de

**Oleksandra Vsesviatska**

oleksandra.vsesviatska@fiz-karlsruhe.de